Received: 1 April 2020 | Revised: 11 January 2021 | Accepted: 28 January 2021

DOI: 10.1111/mec.15837

ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

Detecting genetic signals of selection in heavily bottlenecked reindeer populations by comparing parallel founder events

Menno J. de Jong | Fiona Lovatt | A. Rus Hoelzel

Department of Biosciences, Durham University, Durham, UK

Correspondence
A. Rus Hoelzel, Department of Biosciences, Durham University, Durham, UK.
Email: a.r.hoelzel@durham.ac.uk

Abstract

Founder populations are of special interest to both evolutionary and conservation biologists, but the detection of genetic signals of selection in these populations is challenging due to their demographic history. Geographically separated founder populations likely to have been subjected to similar selection pressures provide an ideal but rare opportunity to overcome these challenges. Here we take advantage of such a situation generated when small, isolated founder populations of reindeer were established on the island of South Georgia, and using this system we look for empirical evidence of selection overcoming strong genetic drift. We generated a 70 k ddRADseq single nucleotide polymorphism database for the two parallel reindeer founder populations and screened for signatures of soft sweeps. We find evidence for a genomic region under selection shared among the two populations, and support our findings with Wright–Fisher model simulations to assess the power and specificity of interpopulation selection scans—namely Bayescan, OutFLANK, PCadapt and a newly developed scan called Genome Wide Differentiation Scan (GWDS)—in the context of pairwise source–founder comparisons. Our simulations indicate that loci under selection in small founder populations are most probably detected by GWDS, and strengthen the hypothesis that the outlier region represents a true locus under selection. We explore possible, relevant functional roles for genes in linkage with the detected outlier loci.

KEYWORDS
adaptation, bottlenecks, natural selection, reindeer, simulation

1 | INTRODUCTION

One of the major current challenges of population geneticists is to discriminate loci under selection from the backdrop of neutral genetic variation (Ahrens et al., 2018; Beaumont, 2005; Oleksyk et al., 2010; Weigand & Lee, 2018). For founder populations, loci under selection are especially hard to detect due to the effects of strong genetic drift. Genetic drift during and following a founder bottleneck spreads out the backdrop of neutral variation, obscuring the typically weak signals of incomplete selective sweeps (Hermisson & Pennings, 2005), and elevating the false negative and positive rates from selection scans. At the same time, founder populations are of special interest to both conservation biologists—due to the impact on biodiversity (e.g., Allendorf & Lundquist, 2003)—and evolutionary biologists due to their potential for promoting rapid evolution (e.g., Templeton, 2008).

Although empirical evidence for adaptation to novel environmental conditions on short, observable timescales has accumulated in recent decades (see Carroll et al., 2007; Endler, 1986; Hendry & Kinnison, 1999; Reznick & Ghalambor, 2001; Schoener, 2011 for reviews on contemporary evolution; and see Bradshaw & Holzapfel, 2010; Brakefield & de Jong, 2011; Hof et al., 2016; Johnston & Selander, 1964; Karell et al., 2011; and Lamichhaney et al., 2015 for case studies), evidence for adaptation specifically in founder populations has so far remained elusive (see Colautti & Lau, 2015; Vandepitte et al., 2014). The rarity of empirical evidence for selection in founder populations will in part reflect adaptive constraints of founder populations (Willi et al., 2006), but also the difficulty of detecting the (genetic) signatures of selection within founder populations (see Poh et al., 2014).

A unique but rare opportunity to overcome the challenges associated with selection analysis in founder populations arises when two or more sister populations (i.e., populations deriving from the same ancestral population) are independently founded in environmentally similar sites (Lee & Coop, 2019). We capitalized on such a system by searching for evidence of selection in two parallel founder reindeer populations. These founder populations originated in the early 20th century (1911 and 1925), when two small ($N \leq 10$) herds of reindeer (*Rangifer tarandus*) were shipped from Filefjell, Norway, to two peninsula separated by a glacier on the island of South Georgia in the South Atlantic Ocean (Leader-Williams, 1980). Despite facing an environment which differed from their native grounds, both populations established successfully until their cull in 2013.

As is true for invasive species in general (Allendorf & Lundquist, 2003; Colautti & Lau, 2015), it is not known whether the success of the South Georgia reindeer was aided by adaptation to their novel environment. We reasoned, however, that if the South Georgia reindeer populations did adapt to their novel environment, the two founder populations potentially underwent parallel evolution. Both populations experienced very similar environmental conditions and therefore would have been subjected to similar selective pressures. If during the founder bottleneck they preserved the same adaptive alleles, this could lead to shared genetic signals of selection. The South Georgia reindeer populations were separated by a glacier over which they could not pass (Leader-Williams, 1988, page 43), and therefore shared signals of selection could not have been established through gene flow. Because shared signals of selection are easier to distinguish from the background of neutral variation than adaptive loci selected in single populations, the South Georgia reindeer populations provide a promising study system for the detection of genomic signals of recent selection within founder populations.

We generated a double digest restriction-site associated DNA sequencing (ddRADseq) database for the founder populations as well as their common source population, and searched for genetic signals of selection using interpopulation selection scans (Oleksyk et al., 2010). Despite concerns over the utility of RADseq for selection scans (Lowry et al., 2016), we reasoned that a large single nucleotide polymorphism (SNP) catalogue would provide sufficient resolution due to the extensive linkage disequilibrium expected in the bottlenecked founder populations. We made use of published selection scan methods (i.e., Bayescan, OutFLANK and PCadapt) together with a custom-built tool which we named Genome Wide Differentiation Scan (GWDS). We evaluated our empirical findings by running simulations using a Wright–Fisher model. The main purpose of these simulations was to estimate the probability that the loci marked as outliers by our whole genome selection scans were true loci under selection rather than false positives. We did so by assessing the power and specificity of selection scans, including GWDS, in the context of founder populations, and specifically in the context of the demographic history of our study populations. In summary, we use Wright–Fisher model simulations to test the hypothesis that positive selection events can be detected despite strong genetic drift in small, severely bottlenecked founder populations, and search for empirical, genetic evidence for positive selection in two recently established founder populations.

2 | MATERIALS AND METHODS

2.1 | Library construction

We selected 120 reindeer samples including some from an existing DNA archive (Table S1, Lovatt & Hoelzel, 2014), evenly divided over both South Georgia founder populations and their Norwegian source population. DNA samples were selected based on Qubit quantification scores and molecular weight of the DNA assessed by gel electrophoresis. We constructed two sequencing libraries each of 60 samples following the ddRADseq protocol (Peterson et al., 2012).

Following *in silico* simulations with the R package SimRAD (Lepais & Weir, 2014), we decided to use a 6-bp cutter (*Hind*III: AAGCTT) and a 4-bp cutter (*Msp*I: CCGG), with a fragment size selection window of 250 bp width (by including all fragments with a length of 275–525 bp, excluding the adapters), targetting 120,000 loci with an average read depth of 30. By multiplying this expected number of loci against their average length (250 bp), as well as with a conservative estimate for nucleotide diversity ($\pi = 0.0005$), and with an approximation for the harmonic number of Watterson's estimator, we estimated that this size selection window would yield ~50,000 SNPs with a minor allele frequency (MAF) ≥0.05.

The actual size selection was executed with a Sage Science PippinPrep machine. The Phusion High-Fidelity kit was used for a 13-cycle polymerase chain reaction (PCR) (denaturation step: 62°C for 20 s; annealing step: 72°C for 45 s; extension step: 72°C for 5 min). Libraries were paired-end sequenced on an Illumina HiSeq_2500 (version 4 chemistry) machine.

2.2 | SNP calling

Reads were trimmed to 110 bp and demultiplexed and filtered using STACKS1.35 (Catchen et al., 2013). Unpaired reads were discarded. Paired reads were aligned using the very-sensitive mode of

BOWTIE2.2.5 (Langmead & Salzberg, 2012), against both the reindeer (*Rangifer tarandus*) genome (Li et al., 2017) as well as the cow (*Bos taurus*) genome (Zimin et al., 2009)—cow being at the time the species closest to reindeer with a genome assembly up to the chromosome level. SAMTOOLS version 1.3.3 (Li et al., 2009) was used to filter out reads which aligned disconcordantly, reads with a mapping quality below 20, as well as reads which aligned to more than one location in the genome.

SNPs were called using the STACKS refmap pipeline with default settings. Loci for which at least 30% of all individuals had a read depth below 8 were removed. We accepted multiple SNPs per read (i.e., we did not set the –write-single-SNPs flag when running the "populations"-command), as we opted to optionally "thin" our data set downstream. The refmap pipeline was initially performed against the cow genome using STACKS version 1.35. After release of the reindeer genome assembly (Li et al., 2017), we used STACKS version 2.2 for SNP calling on the reindeer alignment data set which had by then become available. Unless otherwise stated, the results presented in this paper are based on the SNP data generated with STACKS 2.2.

PGDSPIDER (Lischer & Excoffier, 2012) and PLINK version 1.90 (Purcell et al., 2007) were used to convert the output from genepop or vcf format to a genlight object, supported by the R package Adegenet (Jombart, 2008; Jombart & Ahmed, 2011).

2.3 | SNP and sample filtering

All samples with more than 10% missing data were removed (Table S1). We discarded SNPs which met any of the following criteria: (i) >15% missing data (after removal of low-quality individuals); (ii) minor allele count (MAC) = 1; (iii) excessive heterozygosity excess (He > [2pq + ½q]) (Figure S1A); and (iv) excessive read depth (by fitting a normal distribution to the lowest 99.5% of observed depth values, and subsequently calculating a Bonferroni-corrected threshold value; Figure S1B). We also filtered out a few SNPs which mapped to the same site of the reindeer genome and yet belonged to different STACKS loci. Selection analyses were performed on this filtered data set of 56,079 SNPs. Structure and diversity analyses were performed on a further thinned data set of 33,311 SNPs, which contained at maximum one SNP per 500-bp region (designed to limit SNPs to one per paired-end read).

2.4 | Structure and diversity analyses

For optimization of SNP filter settings, we used a range of SNP data subsets depending on the analysis. For population genetic analyses, we avoided linkage disequilibrium bias by using a filtered and thinned data set. Linkage disequilibrium (LD) analysis was executed on reduced data sets excluding SNPs with MAC < 5, because pairs of SNPs with low number of minor alleles are likely to score high LD-values due to chance effects. For selection analyses, we used a filtered, nonthinned data set, in order to screen the genomes with the highest possible resolution. All but a few analyses were carried out using SNP data derived from alignment to the reindeer genome. The exceptions were genome-wide sliding window genetic differentiation (F_{ST}) and diversity (Tajima's D) analyses, for which we used a filtered, nonthinned data set derived from alignment to the cow genome, as contrary to the reindeer genome the cow genome was assembled to the chromosome level.

All population structure analyses (principal components analysis [PCA], discriminant analysis of principal components [DAPC], admixture analyses) were executed in R, using functions implemented in the Adegenet, Ape (Paradis et al., 2004; Paradis & Schliep, 2018) and LEA (Frichot & François, 2015) packages. For DAPC (run in Adegenet) we set the number of PCs to one third the number of individuals, the number of clusters equal to the number of populations in our data set (three), and the number of discriminant functions to three.

For admixture analysis in LEA we set *K* (putative number of populations) to 2–6, alpha to 10, tolerance to 0.00001 and the number of iterations to 200. To quantify population differentiation we calculated Nei's D (Nei, 1972) using a function implemented in STAMPP (Pembleton et al., 2013), as well as Weir & Cockerham's, 1984 F_{ST} (Weir & Cockerham, 1984). We assessed genetic diversity by generating site frequency spectra, MAF histograms and estimates of sample genome-wide heterozygosity.

To estimate sample genome-wide heterozygosity, we used the formula: $He_{genome} = (n_H/n_{ind} * n_{snps})/n_{total}$, in which n_H denotes the number of heterozygous sites observed for individual i within the SNP data set, n_{ind} denotes the number of nonmissing data points for individual i within the SNP data set, n_{snps} denotes the total number of SNPs in the SNP data set excluding SNPs with excessive heterozygosity (Figure S1), and n_{total} denotes the total number of sequenced sites (i.e., sum of total length of both monomorphic and polymorphic STACKS loci). This provides an estimate of the proportion of heterozygous sites across all sequenced sites, which is a proxy for genome-wide heterozygosity. An illustration of this calculation is provided in the supporting methods (M1).

LD analyses were executed by calculating squared correlation coefficient estimates for unphased data using the software PLINK. We generated LD estimates for all SNP pairs occurring on the same contig at a maximum of 5 Mb apart. Contemporary gene flow was estimated using BayesAss3-SNPs (Mussmann et al., 2019), using the SNP data set generated with STACKS 1.35. The number of iterations was set to 1,000,000, burn-in to 100,000, seed to 10 and delta values to 0.1.

2.5 | Selection analyses

For selection analyses we undertook pairwise comparisons of founder population to source population for each founder population, as well as a "pooled" approach. Our pooled approach assumes that because the two founder populations have been introduced into highly similar environments, they probably experienced shared selective pressures. This allows us to increase the power of our

selection detection analyses when looking for shared signals, because shared outliers would be unlikely to occur by chance alone. In the pooled approach we pooled the data of both founder populations and executed selection scans by contrasting the source population to the pooled founder populations (i.e., "Norway vs. Busen and Barff"). The signal for shared positive selection at a given locus is thereby reinforced whereas the background distribution of neutral alleles is deflated, and hence detection of positively selected alleles is facilitated.

To identify positively selected loci, we used a custom-built tool (GWDS, discussed below) as well as three published selection scans: Bayescan (Foll & Gaggiotti, 2008), OutFLANK (Whitlock & Lotterhos, 2014, 2015) and PCadapt (Duforet-Frebourg et al., 2014; Luu et al., 2017). Bayescan's false discovery rate (FDR) was set to 0.01, as our simulations indicated that higher FDRs decreased specificity without greatly increasing power (Figure S2). We focused on outlier loci with positive alpha scores, which are indicative of positive/diversifying selection (in contrast to negative values, which are indicative of balancing/purifying selection). OutFLANK outliers were scored based on Holm-corrected *p*-values rather than on *q*-values (default setting). PCadapt outliers were scored based on Bonferroni corrected *p*-values, with *K* set to 2. For both the empirical and the simulated data sets, and for both the pairwise and the pooled approach, we ran PCadapt with $K = 2$.

We visually assessed the putative outlier loci by comparing the locus-specific H_E and F_{ST} estimates as formulated by Cockerham and Weir (1987) of putative outliers to those of remaining loci in an H_E-F_{ST} plot (Beaumont & Nichols, 1996). Additional details of the H_E and F_{ST} calculation are provided in the Methods (S2).

2.6 | Selection scan "GWDS"

To assist our selection analyses, we developed a new selection scan called Genome Wide Differentiation Scan (GWDS), which we implemented in the R package "SambaR" (https://github.com/mennodejong1986/SambaR; de Jong et al. 2021) and which searches for loci under positive/diversifying selection on an SNP by SNP basis.

GWDS can be classified as a population differentiation-based method (Oleksyk et al., 2010; Weigand & Leese, 2018) or outlier analysis test (Ahrens et al., 2018), and is most similar to the selection scan OutFLANK. However, unlike OutFLANK and most other inter-population outlier scans, GWDS does not quantify population differentiation using an F_{ST} metric, but instead uses Fisher exact tests, the same metric used by genome-wide association scans (GWAS).

GWDS differs from GWAS in two ways. First, GWDS searches for locus-specific associations between allele frequencies and population divisions, rather than for locus-specific associations between allele frequencies and phenotypic traits. This could concern pairwise comparisons between population pairs, or sets of multiple populations subjected to contrasting environmental pressures.

Allele frequency differences are scored as *p*-values generated by Fisher exact tests executed on contingency tables of allele counts. These *p*-values, here referred to as p_F-values, are calculated using R's built-in fisher.test function, which outputs p_F-values which are up to four decimal places identical to *p*-values outputted by PLINK's (Chang et al., 2015; Purcell et al., 2007) Fisher's exact test.

The second difference between GWDS and GWAS—as well as between GWDS and methods applied in Hendrickson (2013), Cammen et al. (2015) and Shultz et al., 2016)—is in the method of outlier detection. GWDS considers the p_F-values in themselves to be uninformative, as those values depend on sample size, the demographic history of the populations, and the relatedness and connectiveness of the populations. Instead, GWDS searches for loci with p_F-values which stand out from the overall distribution of p_F-values. GWDS does so by fitting an exponential curve to the negative natural logarithms of the p_F-values, and by using this probability distribution to assign new *p*-values to each SNP, here denoted as p_{gwds}-values. These p_{gwds}-values indicate the probability of observing a locus given the observed null distribution. To correct for multiple testing, the p_{gwds}-values are adjusted using either the Bonferroni (default), Holm or Benjamini–Hochberg correction methods. SNPs with an adjusted p_{gwds}-value below .05 (or any other user-defined significance threshold) are marked by GWDS as outliers. An illustrated step-by-step description of GWDS is included in the supporting methods (M3).

The logic behind GWDS is that whereas allele frequencies of neutral loci differ randomly among populations, selection will temporarily cause the allele frequencies of positively selected loci to differ more strongly. This reasoning especially holds if compared populations are exchanging migrants or alternatively if they are sister populations (i.e., derived from the same ancestral population), because their allele frequencies are initially correlated. In the absence of gene flow, a requirement is that the time to the most common recent ancestor of the sister populations is considerably less than $4Ne$ generations, as greater split times will result in differential loss or fixation of alleles through drift.

By inferring the null distribution directly from the empirical data rather than attempting to model the expected null distribution, GWDS, like OutFLANK, circumvents the risk of violating underlying assumptions about demography, such as population hierarchy, changes in population size through time, and unequal levels of gene flow (Funk et al., 2016; Luu et al., 2017; Weigand & Leese, 2018). The main assumption of GWDS is that the distribution of the negative natural logarithms of observed p_F-values can be described by an exponential curve. This assumption holds if and only if both sister populations exchange migrants or split less than $4Ne$ generations ago. In the absence of gene flow, longer split times will result in a bimodal distribution, with the modi reflecting differential fixation and loss of alleles due to drift. GWDS tests the goodness of fit between observed data and fitted curve using a chi-squared test and will issue a warning in the case of a poor fit. Another assumption behind GWDS is that the vast majority of SNPs will be neutral, and that the observed mean will not be inflated by a few loci under selection. This assumption can be violated when applying GWDS to dense SNP catalogues in which multiple SNPs are linked to the same

selective sweep, resulting in a too conservative outlier threshold. GWDS therefore offers users the option to fit the exponential curve to a thinned data set which contains at most one SNP per genomic region of 1 Mb (or any other user-defined size).

We tested the power and specificity of GWDS under various demographic scenarios of recently established founder populations and compared these test scores to the power and specificity of two popular selection scans which also run in R, OutFLANK and PCadapt (see section Wright–Fisher model simulation analyses).

2.7 | Wright–Fisher model simulator

We simulated data sets of founder and source populations using custom R functions based on a Wright–Fisher model. The demographic scenario consisted of a source population with a constant *Ne* of 1,000 individuals from which a founder population buds off at t_0. Both the source and the founder population are subsequently allowed to drift for a certain number of generations. The source and the founder population do not exchange migrants (i.e., no gene flow).

The simulation tool simulates changes in allele frequencies of standing variation through generations on a SNP by SNP basis. It does not incorporate new mutations, and all SNPs are assumed to be unlinked. The distribution of allele frequencies in the source population were modelled after the observed distribution in the empirical data set (see Figure S3). Founder events and genetic drift subsequent and prior to the founder event were simulated with the rbinom function, which outputs the number of successes (number of allele copies in next generation) given a sample size (2*Ne*) and a success probability (allele frequency in current generation).

Selection was simulated in two ways. For selection coefficients of $s > 0.05$, we simulated selection as a continuous process by multiplying each generation by the rbinom output with the selection factor $(1 + s)$. For $s \leq 0.05$, we opted for a different approach because the effect of selection was counteracted by rounding. Here we simulated a selective event as a doubling of the number of adaptive alleles, with a probability of occurrence of s per generation (Method S4). Neutral loci were defined as SNPs for which allele frequencies were affected by drift only. Loci under selection were defined as SNPs for which allele frequencies were affected not only by drift, but in the case of founder populations also by positive selection, either directly or indirectly (through complete linkage). The envisaged evolutionary scenario was that these alleles had been (nearly) neutral in the source population, occurring in low frequencies in a mutation–drift equilibrium. Following environmental change due to the colonization of a new habitat, these previously neutral and rare alleles became positively selected in the founder populations. It was assumed that selection results in allele frequency changes but not in population size changes.

The outputs of our simulations were allele frequencies/counts of a source and two founder populations, both directly following the founder event (t_0) and after a certain number of generations (t_{gen}). To incorporate observer error related to limited sampled sizes (i.e., deviation between population allele frequencies and observed allele frequencies), we generated sample allele frequencies using the rbinom function, with the number of successes representing the number of allele copies in the genotyped individuals given, a sample size equalling 30 individuals, and with success probability being defined as the population allele frequency. Sampled t_{gen} output vectors served as simulated input for selection scans.

We validated our simulator by comparing three simulation output scores with theoretical expectations: (i) the proportion of retained variation directly after a founder event; (ii) the fixation probability of neutral and adaptive alleles; and (iii) time to fixation.

2.8 | Wright–Fisher model simulation analyses

We used our Wright–Fisher model simulator to assess the performances of GWDS in comparison to PCadapt and OutFLANK. Bayescan, which runs on Linux rather than in R and has relatively long computation times, was excluded from this part of the analysis. For our purposes the three-way comparison was considered sufficient. Both Bayescan and OutFLANK are F_{ST} outlier tests, whereas PCadapt represents a distinct method (based instead on PCA). For each test—GWDS, OutFLANK and PCadapt—we calculated the false positive rate (1 − specificity) as the number of neutral SNPs marked as outliers divided by the number of neutral SNPs with MAF > 0 at t_0. Similarly, for each test we calculated the power (1 − false negative rate) as the number of selected SNPs marked as outliers divided by the number of selected SNPs with MAF > 0 at t_0. In other words, adaptive loci which were lost during the founder bottleneck event were excluded from the power and specificity calculations.

FDR estimates, the proportions of false positives in the outlier set, were calculated based on power and specificity estimates inferred using the Bonferroni multiple testing correction method, and were calculated for various proportions of adaptive SNPs in SNP data sets (which is an unknown parameter, here defined as *p*_adaptive), ranging from 10% to 0.01%. The exact formula was: (FP*(1 − *p*_adaptive))/(FP*(1 − *p*_adaptive) + ((1 − FN)**p*_adaptive)), in which FP denotes false positive rate (inverse of specificity) and FN denotes false negative rate (inverse of power).

We evaluated the performance of the three selection scans for a range of demographic scenarios, which included all combinations of selection coefficients ranging between 0 and 0.2 (step size 0.025) and constant effective founder population sizes of 10, 20, 30, 50, 100, 200, 300, 500 and 1,000. The number of founders was set equal to the effective population size, and the number of generations of the founder population (starting at the founder event) was set to 20. In addition, we ran simulations for a demographic scenario specific to our South Georgia reindeer populations, defined by a bottleneck event of 10 founders and followed by a constant effective population size of 50 individuals during 20 subsequent generations, for a total of 60,000 SNPs.

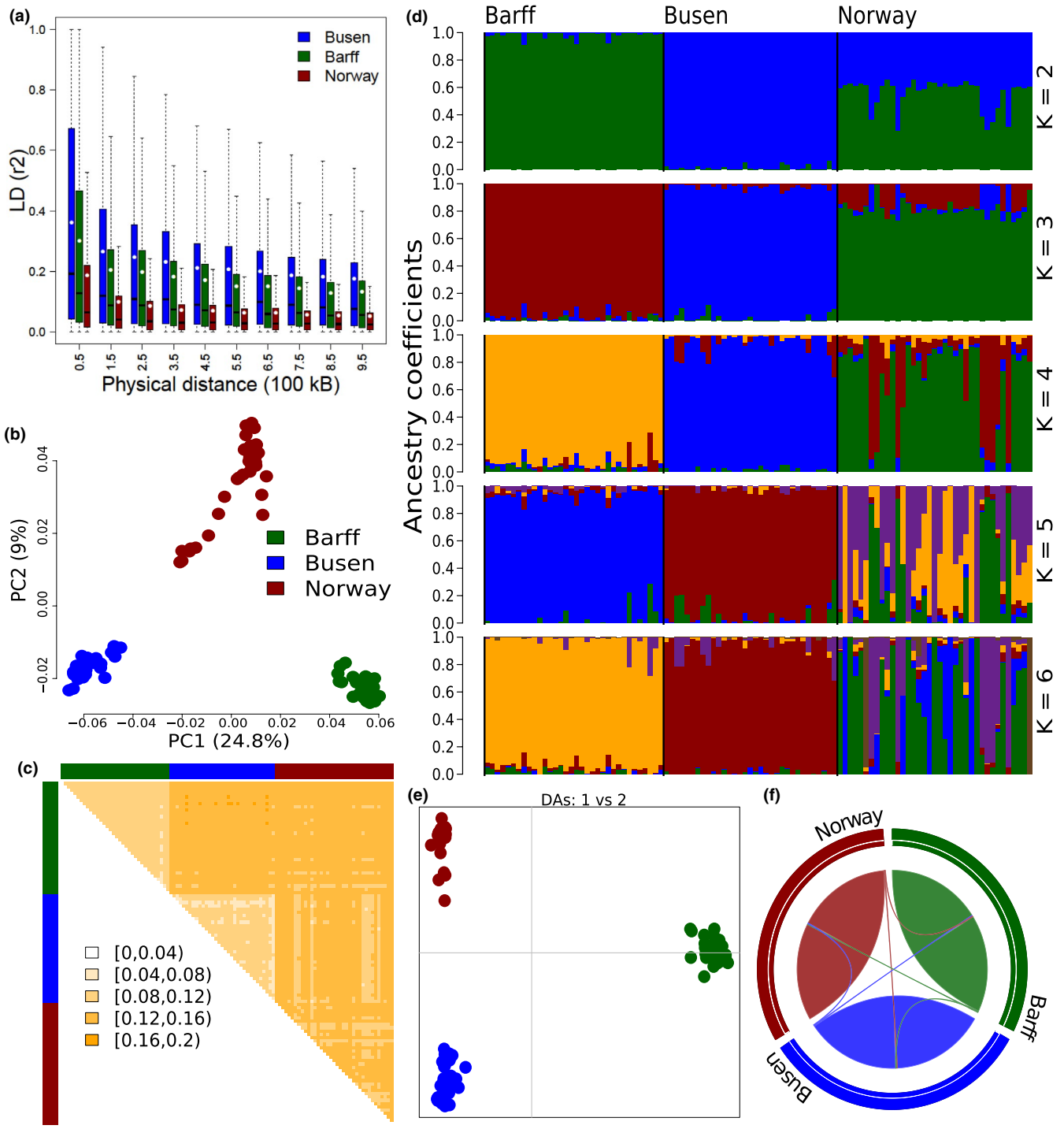

FIGURE 1 Genetic clustering and linkage disequilibrium (LD) analyses of reindeer samples belonging to both South Georgia populations (Busen and Barff) and their Norwegian source population. Colour coding (except for D): blue = Busen, green = Barff, red = Norway. (a) Boxplots of LD estimates (squared Pearson correlation coefficients based on genotype scores of SNP pairs, calculated by the software PLINK) per physical distance class (100kb bins). White dots indicate mean values. (b) Principal coordinates analysis based on Nei's genetic distance. (c) Nei's genetic distance between samples. Distance is proportional to colour intensity. (d) LEA admixture analyses for $2 \leq K \leq 6$, with random colour coding. (e) Discriminant analysis of principal components, without prior population information, with $K = 3$ and a number of PCs which explains 80% of the cumulative variance. (f) Migration rates between the three populations, as inferred by Bayesass3-SNPs [Colour figure can be viewed at wileyonlinelibrary.com]

GWDS and OutFLANK were instructed to calculate the neutral distributions based on the neutral loci only. Because this option is not available for PCadapt, and so that PCadapt could reliably obtain a neutral distribution, we set the proportion of adaptive alleles to a maximum of 0.1. We have evaluated the performance of PCadapt using varying proportions of adaptive SNPs, and found that PCadapt behaves normally (no elevated false positive or false negative rates) on data sets with a maximum of 10% adaptive SNPs.

2.9 | Gene identification

Genes and other genomic features close to outlier SNPs were identified based on both a cow genome annotation (Bovine Genome Sequencing and Analysis Consortium et al., 2009) and a reindeer genome annotation (Li et al., 2017), using the software BEDTOOLS version 2.26.0 (Quinlan & Hall, 2010). Genes were considered putatively linked to an outlier SNP if they were within 150 kb distance from the outlier SNP, and if no nonoutlier SNP was present in between. LD calculations indicated that this range spanned most of the SNP pairs in high linkage (Figure 1a).

3 | RESULTS

3.1 | Structure and diversity analyses

We used ddRADseq (Peterson et al., 2012) to generate an ~70 k SNP database, located on ~85 k STACKS loci (De Jong et al., 2020a, b2020). After filtering, we retained ~56 k SNPs and ≥30 individuals per population (Norwegian source population: 37; founder population "Busen": 33; founder population "Barff": 34; Tables S1–S3). Population structure analyses (i.e., PCA, DAPC, admixture analyses; based on a set of ~33 k SNPs thinned to minimize LD) verified the existence of three distinct clusters, and therefore supported the assumption that the two founder populations were geographically isolated (Figure 1; Figures S4 and S5). Absence of gene flow was furthermore confirmed with the software BayesAss3-SNPs (Figure 1; Table S4). As expected based on the stochastic sampling of alleles during and following the founder events, and as previously illustrated for these populations (Lovatt & Hoelzel, 2014), each founder population was differentiated from the other and both were differentiated from the source population (based on ordination, assignment, genetic distance and gene flow analyses; see Figure 1). Weir and Cockerham's (1984) F_{ST}-values for the three pairwise population comparisons were 0.197 for Barff vs. Busen, 0.087 for Barff vs. Norway, and 0.112 for Busen vs. Norway.

Population-specific genetic diversity estimates showed strong signatures of recent bottleneck events, with both founder populations displaying site frequency spectra (SFS) typical for bottlenecked populations: reduced nucleotide diversity coupled with high proportions of common SNPs, testifying that many alleles, mostly of low frequency, were lost during and/or after the founder bottlenecks (Figures S6 and S7, Table S5).

LD analyses confirmed the presence of extended LD in both founder populations (Figure 1a), facilitating our genome-wide selection scans. The observed decay of mean LD-values with physical distance roughly corresponded with expected values given recent effective population sizes of respectively 50, 80 and 1,000 for Busen, Barff and Norway (Figure S6F,G, Hill & Weir, 1988; McVean & Cardin, 2005: Kawakami et al., 2014).

Estimates of genome-wide proportions of segregating sites were 0.34%, 0.39% and 0.69% respectively for Busen, Barff and Norway (Figures S6C, S7E and S8, Table S5). The observed differences in the proportion of segregating sites between populations could not be explained by variation in levels of missing data (Figure S7A), and only marginally by the difference in number of retained individuals per population (Figure S7B). The equation $1 - (1 - \text{MAF})^{2N\_founders}$ can be used to calculate expected proportions of segregating sites directly following the founder bottleneck. Given that the average MAF within Norway was 0.17, and knowing that the number of founders of the Busen and Barff populations were respectively seven and 10 individuals (Leader-Williams, 1988), the proportion of segregating sites directly following the bottleneck will have been around respectively 0.64% and 0.67%, much higher than the observed 0.34% and 0.39%. The implication is that the majority of genetic variation was lost due to genetic drift during subsequent generations rather than during the founder event itself.

3.2 | Selection analyses

When screening our filtered SNP data set (~56 k SNPs) for loci under selection, we contrasted the pairwise comparison of each founder population to its source population with a "pooled" approach that took advantage of the increased power availed when selection was affecting a locus in each founder population in the same way. Note that only loci that are evolving in parallel (adapting in the same way to highly similar environments) in the two founder populations could be detected by the pooled approach (see Methods for further details). Our expectation for the pairwise approach was that there would be insufficient power to reliably detect true outliers, due to strong drift, as found in earlier studies (e.g., Poh et al., 2014).

The number of outliers identified by Bayescan (Figure S9A), GWDS (Figure S9B), OutFLANK (Figure S9C) and PCadapt for the pooled approach were respectively 10, three, five and 15 (Figure 2). Overlap between the sets of outliers identified by different scans was restricted to two outliers marked by all selection scans (Figure 2c; Figure S9D,E) and one outlier marked by both GWDS and OutFLANK. All other SNPs marked by Bayescan were fixed in both founder populations for the dominant allele in the source population (Figure S9G). All other SNPs marked by PCadapt differed strongly between a founder population and the other two populations, rather than between the source and both founder populations (Figure S9G). The two outlier SNPs called by all four selection scans had GWDS and OutFLANK $\log_{10}$-converted p-values above respectively 6.62 and 7.16, compared to a Bonferroni threshold of 6.05. The Bayes Factors output by Bayescan were 126 and 1,660. According to alignments to both the cow genome and the reindeer genome, the two outliers were adjacent SNPs 80–85 kb apart. They mapped to a genomic region of cow chromosome 25 that displays a signature indicative of positive selection in sister populations: F_{ST} peaks for each source–founder comparison, and an F_{ST} valley for the founder–founder comparison (Figure 2b; Figures S10 and S11). As expected for a soft and incomplete selective sweep, sliding window Tajima's D analyses did not reveal a

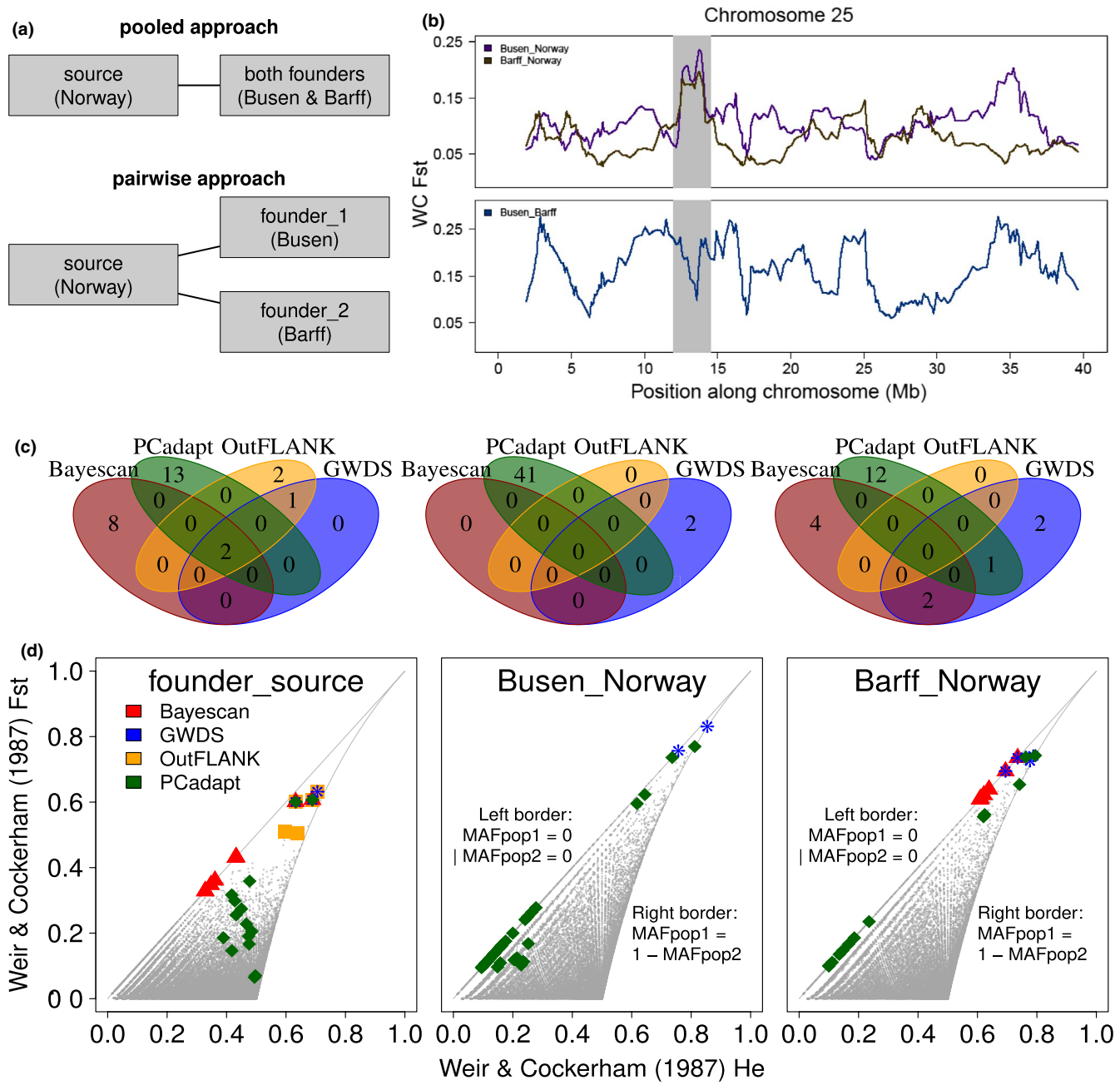

FIGURE 2 Selection analysis. (a) Conceptual model of the two approaches used when running selection analysis. (b) Window F_{ST} values around position 14 Mb on chromosome 25, the location of the two adjacent outlier SNPs. (c) Venn diagrams of outlier sets outputted by the selection scans Bayescan, GWDS, PCadapt and OutFLANK for both the pooled approach (founder-source) as well as both comparisons of the pairwise approach. (d) Fdist plots showing the location of outlier SNPs output by the selection scans Bayescan, GWDS, PCadapt and OutFLANK for the pooled approach as well as both comparisons of the pairwise approach [Colour figure can be viewed at wileyonlinelibrary.com]

signal of selection for this genomic region (Figure S12), although the power is relatively low using RADseq data for this analysis. The population-specific MAFs (with the minor allele defined respective to the metapopulation) of the most confidently marked outlier SNP equalled 0.05, 0 and 0.64 for respectively Busen, Barff and the Norwegian source population (Figure S9G). The adjacent SNP had MAFs of 0.11, 0.04 and 0.72 (Figure S9G). Hence, the two adjacent outlier SNPs show a consistent signal of positive selection on an allele which was the minor allele in the source population.

None of the outliers identified in the pooled approach were identified by any of the selection scans for pairwise population comparisons (Figure S9F). Also, none of the four tests showed overlap between the outliers identified in the pairwise comparisons (i.e., outliers of the Busen–Norway comparisons differed from outliers of the Barff–Norway comparisons). Overlap between selection scan outlier sets per pairwise comparison was restricted to the Barff–Norway comparison, with one SNP marked as outlier by both GWDS and Bayescan, and two SNPs marked by both GWDS and Bayescan (Figure 2d; Figure S9E,G).

H_E–F_{ST} scores of putative outlier loci clustered to some extent by selection scan (Figure 2d). The outlier sets of both PCadapt and Bayescan contained predominantly SNPs which were fixed in one of either population (Figure S9G). For pairwise comparisons (i.e., Barff–Norway and Busen–Norway), the opportunity for putative outlier loci to stand out from neutral loci was limited because the overall H_E–F_{ST} distribution filled the entire spectrum of possible H_E–F_{ST} values for pairwise population comparisons. The overall distribution of H_E–F_{ST} estimates was less inflated for the pooled data set compared to either pairwise data sets (Figure 2d), increasing the opportunity for loci under selection to stand out from the neutral distribution and hence to be detectable by selection scans.

3.3 | Wright–Fisher model simulations

To better understand the observed inconsistencies, we estimated the power and specificity of the selection scans Bayescan, GWDS, OutFLANK and PCadapt on simulated data generated with a custom-built Wright–Fisher model simulator. This simulation tool was specifically designed to simulate unlinked neutral and adaptive allele frequencies in founder and source populations following a founder event.

We validated the model by comparing simulation results (i.e., proportion of retained alleles and fixation probability and time) with established equations from theoretical population genetics (see Methods; Figure S13). Consistent with expectations, the observed proportion of retained variation depended on the number of founders ($N_{founders}$) and allele frequencies in the source population as follows: $1 - (1 - MAF)^{2N\_founders}$ (Figure S13A). Fixation probabilities approximated $(1 - e^{-2Ne \cdot sp})/(1 - e^{-2Ne \cdot s})$ (Figure S13B, Kimura, 1962), which for nearly neutral alleles corresponded to the mean frequency of alleles directly following the founder event, as expected for neutral alleles ($s = 0$) (Figure S13B). We also confirmed that fixation times were less than $4Ne$ generations for neutral alleles (Figure S13C) and less than $(2/s) \cdot \ln(2Ne)$ generations for adaptive alleles (Figure S13D, Kimura & Ohta, 1969). This is consistent with expectations, because the fixation time of standing variation should fall below the fixation time of new mutations.

Our simulations indicated that selection scans generally fail to detect the majority of positively selected loci in founder populations which are founded recently (<20 generations ago) and which are small to moderate in size ($Ne < 100$) (Figure 3; Figure S14, Table S6). Specificity scores are generally close to 1 (>99.99% in the case of GWDS and OutFLANK), but can nevertheless lead to high false discovery rates (e.g., >50%) depending on the proportion of adaptive

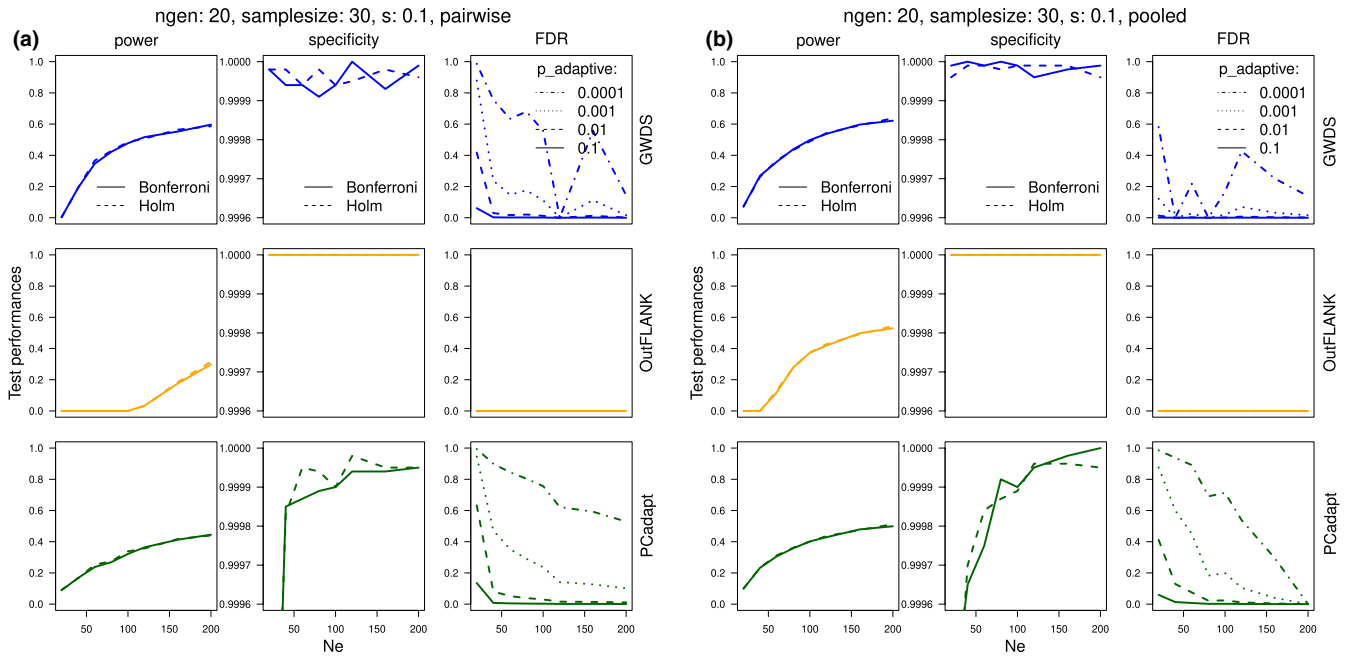

FIGURE 3 False discovery rates (FDR) of selection scans in young founder populations. Power, specificity and FDR estimates of the selection scans GWDS, OutFLANK and PCadapt in recently established founder populations (population age of 20 generations) given a sample size of 30 individuals per population, a selection coefficient s of strength 0.1, various constant effective population sizes (Ne) without founder bottleneck, and using either the Bonferroni or Holm multiple test correction method. Power estimates give the inverse of the false negative rate (1 − FN), that is the proportion of alleles under positive selection that are correctly marked by selection scans as outliers. Specificity estimates give the inverse of the false positive rate (1 − FP), that is the proportion of neutral alleles that are not marked by selection scans as outliers. The power and specificity scores are based on simulations with 90,000 neutral SNPs and 10,000 adaptive SNPs. FDR estimates, the proportions of false positives in the outlier set, are based on the Bonferroni power and specificity estimates and are calculated for various proportions of adaptive SNPs (p_adaptive), ranging from 10% to 0.01%, using the formula: (FP*(1 − p_adaptive))/(FP*(1 − p_adaptive) + ((1 − FN)*p_adaptive)). Left: pairwise approach. Right: pooled approach [Colour figure can be viewed at wileyonlinelibrary.com]

SNPs in the input data set, an unknown parameter (Figure 3; Figure S14, Table S6). We found that GWDS has higher power than PCadapt and OutFLANK for scenarios involving relatively low effective population sizes (*Ne* <50) or inflated null distributions (as in the case of the pairwise approach; Figure 3; Figure S14), without compromising specificity beyond intolerable levels. The simulations also showed that the Bonferroni and Holm correction methods lead to similar results (Figure 3) and are more conservative than the Benjamini–Hochberg method (Figure S14B).

In small founder populations (e.g., *Ne =20*), in which drift is dominant, the selected loci have a bimodal distribution on the line $H_E = F_{ST}$ (Figure S15), indicating fixation in one of either population (supporting methods 2). The group with low H_E and F_{ST} scores represents mostly loci which were lost in the founder population after the founder event, due to genetic drift overruling the workings of positive selection. The group with high H_E and F_{ST} scores represents mostly loci which due to selection reached fixation in the founder population. The proportion of selected loci belonging to the first group decreases with increasing *Ne* (Figure S15).

Next we modelled the demographic history of the South Georgia reindeer populations as a two-step demographic scenario, consisting of a bottleneck of 10 effective founders for one generation, and a fixed *Ne* of 50 individuals during 20 subsequent generations. For this demographic scenario Bayescan, GWDS, OutFLANK and PCadapt

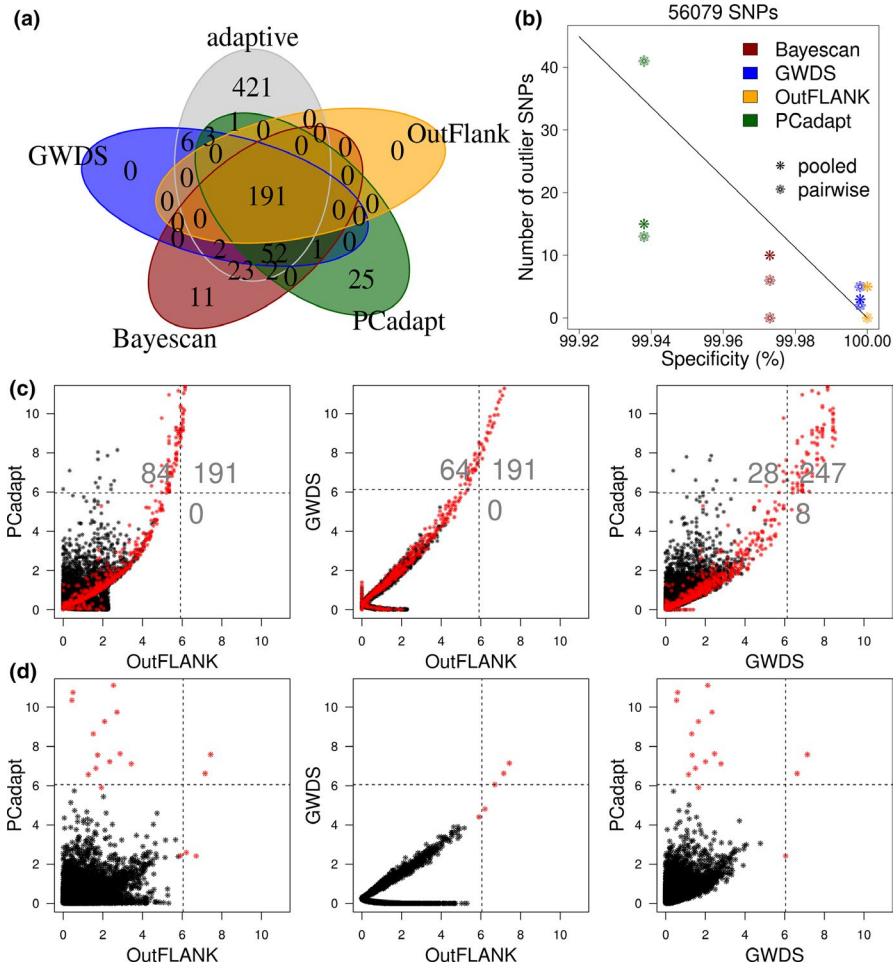

FIGURE 4 Overlap between selection scans. Overlap between outliers scored by the selection scans Bayescan, GWDS, OutFLANK and PCadapt in simulated (a,c) and empirical (d) data sets for the pooled comparison. Simulations are based on 59,000 neutral SNPs and 1,000 adaptive SNPs (s = 0.1), and a two-step demographic scenario meant to reflect historical *Ne* of both South Georgia reindeer populations: a bottleneck of 10 individuals for one generation, and a fixed *Ne* of 50 individuals during 20 subsequent generations. In total, 56,079 neutral and 701 adaptive SNPs were retained in both founder populations during the founder events. (a) Venn diagram showing the simulated overlap between outlier sets and true loci under positive selection. (b) Expected number of false positives (black line), calculated as (1 − specificity)*56,079 SNPs, vs. the number of putative outliers outputted by selection scans for all three comparisons (i.e., Barff vs. Norway, Busen vs. Norway, and Barff and Busen vs. source). Specificity estimates were calculated from simulated data (see Figure 3). (c) Scatterplots comparing selection scan output statistics of simulated neutral (black) and selected (red) loci using the pooled approach. Dashed lines indicate outlier thresholds (i.e., all loci above these thresholds are marked by the selection scan as outliers). Numbers indicate the number of adaptive SNPs per segment. Bayescan p-values (posterior probabilities) cannot be directly compared to the frequentist's p-values of the other three selection scans, and are therefore omitted. (d) As (c), but for empirical rather than simulated data sets. Red dots indicate the SNPs marked by either of the two selection scans [Colour figure can be viewed at wileyonlinelibrary.com]

marked respectively 12, one, zero and 26 out of 59,000 neutral loci as false positives, translating to specificity scores of respectively 99.980%, 99.998%, 100% and 99.956% (Figure 4a; Figure S16). The total number of outlier SNPs marked by the four selection scans in our empirical data sets fits the expected number of false positives based on these specificity scores and the size of our data sets (Figure 4b). The simulations furthermore suggested that (given the demographic scenario described above):

1. nearly all SNPs marked by two (or more) scans are true adaptive loci (Figure 4c);
2. the majority of adaptive loci which are marked as outliers with the pooled approach are not marked as outlier with the pairwise approach, and vice versa;
3. when plotting the GWDS test scores of both independent pairwise comparisons against each other, our three putative loci identified from the pooled approach are positioned in a plot region which according to our simulations holds adaptive loci exclusively.

3.4 | Identification of genomic features in proximity to outlier SNPs

The closest known gene to the nonadjacent outlier SNP is HAO1, which codes for the protein hydroxyacid oxidase. This gene is, however, separated from the outlier SNP by a stretch of 200 kb containing four nonoutlier SNPs, and is therefore unlikely to be of interest (Figure S17).

Alignments to both the reindeer and the cow genome indicated the presence of an exon in between the two adjacent outlier SNPs (Figure S17). This exon is part of a gene coding for myocardin-related transcription factor B, known as both MRTF-B and MKL2. MKL2, short for megakaryoblastic leukaemia 2, is a member of the myocardin family (Selvaraj & Prywes, 2003). This family contains the protein myocardin (MYOCD), the transcription factors A and B (MKL1 and MKL2), and MASTR (Swärd et al., 2016).

4 | DISCUSSION

In this study we capitalized on a seminatural experiment to search for shared signals of selection in two sister reindeer populations which were founded in geographically isolated but environmentally similar habitats a similar number of generations ago. To evaluate our empirical findings, we assessed the power and specificity of selection scans using simulated data sets generated with a custom-built Wright–Fisher model simulator. Our overall aim was to gather empirical evidence that founder populations can start adapting directly following a founder event, even when the bottleneck effect is severe, though of course this does not provide the level of confidence about adaptation that would be possible in a controlled experiment.

4.1 | Genetic evidence for adaptation within South Georgia reindeer populations?

Inconsistency in outlier sets produced by different selection scans, as reported in this study as well as other studies (e.g., figure 2a in Andrew et al., 2018; figure 3a in Chen et al., 2018; figure 5 in Heppenheimer et al., 2018; figure 3 in Fuller et al., 2019, figure 5 in Takeuchi et al., 2019), can result from both false positives and false negatives. Our pooled approach identified three SNPs which were marked as outliers by two or more selection scans. Two of these SNPs were detected by all four methods, whereas the third was identified by GWDS and OutFLANK (Figure 2c). Our simulations indicated that within our data sets false positives are uncommon among SNPs detected by two and especially by three selection scans (i.e., Bayescan, GWDS and PCadapt, Figure 4a), and therefore imply these three SNPs are probably true loci under (linked) selection (Figure 4c,d). Our simulations furthermore show that this conclusion is not contradicted by the fact that these SNPs were not detected by either of the pairwise comparisons (i.e., Busen–Norway and Barff–Norway).

A potential confounding factor which cannot be easily assessed through simulations is the effect of genotyping errors. However, the adjacent positioning of two of the outlier SNPs argues against genotyping errors. These two adjacent SNPs share a congruent signal of selection despite being located on different sequencing reads. The improbability of any pair of unrelated outlier SNPs being adjacent by chance, given the small proportion of outlier SNPs (three out of ~56 k SNPs in total), greatly diminishes the chance that their unusually high F_{ST} values result from sequencing errors.

Concerns have been raised that RADseq-based selection scans might lack the resolution to detect selective sweeps, because the spacing between SNPs might exceed the extent of LD (Lowry et al., 2016). The detection of adjacent outlier SNPs, being in high linkage despite considerable physical distance (80 kb), provides strength to the belief that under certain circumstances low-density SNP catalogues can suffice to detect at least a subset of loci under selection (Catchen et al., 2017). Population bottlenecks complicate, through the workings of genetic drift, the detection of adaptive loci, but the effect cuts both ways. The associated decrease in population recombination rate can in turn facilitate outlier detection. Observed and predicted mean LD values in our two founder populations (Figure 1a; Figure S6F,G), was approximately 0.3–0.4 at 100-kb distance, which corresponds to Pearson correlation coefficients ranging between 0.55 and 0.63. More than 25% of all SNP pairs within 100-kb distance showed LD-scores above 0.5, corresponding to Pearson correlation coefficients of 0.7 or higher.

It is also possible for variation in selection or the recombination rate across the genome to influence the detectability of signals for selection. Recombination rate variation determines the size of a selective sweep, and hence detectability. Therefore, outlier detection is facilitated if positive selection occurs on an allele in a region with low recombination rate. On the other hand, it could be argued that high recombination rates can more easily break down linkage to a potential neighbouring allele under purifying selection, and thereby

allow the establishment of small selective sweeps which are difficult to detect but would not have occurred in regions of low recombination rates.

According to the SNP data set generated with STACKS version 1.35, one of the two loci containing an adjacent outlier SNP contained next to the outlier SNP a neutral SNP. This neutral SNP, 34 bp distant from the outlier SNP, had population-specific MAFs of 0, 0 and 0.01 for respectively Busen, Barff and Norway (source population; Figure S16). Given the improbability of a recombination event at such a short distance (Johnston et al., 2017), the different fates of both alleles can only be explained if these alleles were not in complete linkage in the original founder population. The only copy of the adjacent neutral allele sampled in the source population occurred in an individual which was heterozygous for both the neutral SNP and the outlier SNP, providing little scope to estimate linkage. Furthermore, this particular SNP was not present in the SNP data set generated by STACKS 2.2.

Our simulations assumed an evolutionary scenario in which (nearly) neutral alleles occur originally at low frequency in mutation-drift equilibrium in the source population, and become adaptive and eventually fixed in the founder populations due recolonization to a new environment. The fact that in the source population the alleles of interest have higher frequencies (1 − 0.64 = 0.36, 1 − 0.72 − 0.28) than the average MAFs in the source population (~0.15) might appear in disagreement with this hypothesis. However, it can be argued that low-frequency alleles are less likely to survive both founder bottlenecks. Because parallel selection can occur only on alleles which are present in both founder populations, it is perhaps not surprising that the outlier alleles have relatively high frequencies in the source population.

One alternative evolutionary scenario could be that the selective event has occurred in the source population rather than in the founder populations, due to changes in the source population environment occurring after the founder events. However, the amount of environmental change in the source population is likely to have been far less over this time frame than that experienced during the transfer to the novel environment experienced by the founder groups in South Georgia. The higher frequencies in the source population may also simply be neutral alleles drifting towards fixation in a finite population.

4.2 | Hypothetical phenotypic trait under selection

If the identified outlier region(s) are indeed true loci under positive selection in the founder populations, what were the associated phenotypic traits under selection? Insular populations famously exhibit evolutionary trends in both morphological and behavioural traits (Losos & Ricklefs, 2009). One of these trends, the island rule or Foster's rule, involves changes in body size and predicts dwarfing of big species and gigantism of small species (Foster, 1964; Lomolino et al., 2013; Rozzi & Lomolino, 2017). Cervidae are among the taxonomic groups which are particularly susceptible to insular dwarfing (Lomolino et al., 2013), and reindeer are no exception. Multiple insular populations of reindeer are characterized by reduced leg length, most extremely the Svalbard reindeer (Klein et al., 1987). No evidence exists, however, for decreased leg lengths or decreased body size in South Georgia reindeer (Leader-Williams, 1988).

Rather than being associated with insularity, it is also possible that traits under selection in the South Georgia populations were associated with factors specific for the South Georgia habitat. Environmental differences between South Georgia and the habitat of the Norwegian source population included a higher salinity (sea spray and greater proportion of marine grasses), the absence of predators, a milder climate (although with more heavy winds, Leader-Williams, 1988, p. 36), and dietary changes due to vegetation differences. According to investigations by Leader-Williams (1988), this last category might have led to increased mortality rates among the South Georgia reindeer.

South Georgia reindeer mortalities followed patterns typical for deer, with females dying mostly in late winter and males dying mostly in early winter, after the rut (Leader-Williams, 1988). However, one mortality factor not commonly observed for insular populations, nor in the Norwegian source reindeer population, was dental disease (Leader-Williams, 1988). Symptoms varied from missing to split or broken mandibular premolars and molars, regularly accompanied by mandibular swellings (Leader-Williams, 1982). These mandibular swellings affected 9%–19% of all individuals within both populations (Leader-Williams, 1982, table 1). As mandibular swellings are likely to reduce the efficiency of chewing and therefore energy uptake, they could affect survivability. Indeed, Leader-Williams (Leader-Williams, 1982, table 3) found that 22.9% of over 100 examined carcasses were affected, whereas based on the prevalence in living individuals a proportion of 15.1% was expected. Field observations also suggest that affected individuals coupled their higher mortality rates with lower fecundity (Leader-Williams, 1988, p.177).

Both radiographic and chemical analyses showed severe osteoporosis of mandibles, increasing with age and being more pronounced in individuals with mandibular swellings (Leader-Williams, 1988, p.174). Leader-Williams (1988) hypothesized a scenario in which a combination of overpopulation and limited availability of nutrient-rich vegetation led to mineral deficiencies in the South Georgia reindeer. This caused osteoporosis in mandibulars, and increased susceptibility for tooth damage and tooth loss (Darcey et al., 2013), perhaps aided by increased susceptibility for infections by micro-organisms (Leader-Williams, 1988, p.175). We suggest that the South Georgia reindeer might have possessed heritable variation in susceptibility to mandibular osteoporosis and tooth damage, resulting from the presence of a polymorphism within MKL2 itself or within *cis*-regulatory elements. Although we agree that mineral deficiencies in the newly colonized environment could explain the sudden manifestation of a previously unseen condition, we also note that genomic stress resulting from bottlenecks can impact morphology (Lovatt & Hoelzel, 2011).

MKL2 enables protein kinase activity and ATP binding, and is involved in protein phosphorylation. It is a transcriptional

coactivator of the serum response transcription factor (SRF). SRF controls the expression of muscle-specific genes, and is required for both striated and smooth muscle differentiation (Selvaraj & Prywes, 2003; Swärd et al., 2016). The exact mechanism by which an MKL2 allele could have counteracted mandibular osteoporosis and tooth damage despite mineral deficiencies is unknown, and hypothesized mechanisms are speculative by nature. However, one possible mechanism is that MKL2 variants infer increased resistance to tooth disease by acting upon E-cadherin (Guo et al., 2014). E-cadherin enables calcium ion binding and is involved in cell adhesion. E-cadherin-mediated cell–cell adhesion and signalling plays an essential role in the development and maintenance of healthy epithelial tissues (Guo et al., 2014). Teeth have a mesenchymal as well as an epithelial component, and E-cadherin is thought to regulate odontogenesis (Heymann et al., 2002; Li et al., 2012).

4.3 | Adaptive constraints of founder populations?

Theory predicts that bottlenecks negatively affect the adaptive capacities of founder populations through a reduction of genetic variability (i.e., reduction of adaptive potential, Willi et al., 2006) and a temporal increase of the magnitude of genetic drift.

Our simulations and empirical data indeed indicate severe loss of genetic variation within the South Georgia founder populations. As a result, only a minority of potential adaptive alleles can be expected to have been retained in both South Georgia founder populations. The implication is that most selective events are expected to be private events, limiting the potential for parallel adaptive evolution. Hence, although parallel adaptation can help overcome the difficulties associated with selection analyses (especially so in founder populations), practical limitations need to be taken into consideration. Adding additional populations does not guarantee increased power, because the probability of a shared selective event in all populations decreases with the number of populations investigated. This might explain the findings in earlier studies on parallel populations, which has provided limited evidence for shared selective events (Perrier et al., 2013; Roesti et al., 2014).

With regard to the increased magnitude of genetic drift due to small population size, our simulations confirm that even in the face of strong genetic drift, selection of sufficient strength (e.g., $s = 0.1$) can drive a proportion of adaptive alleles to fixation within a relatively short time frame (e.g., 20 generations) depending on the magnitude of drift (as determined by *Ne*; Figure S14). Of course, if an adaptive allele happened to increase in proportion following stochastic sampling during a founder event, that may favour its chance of fixation, though this will still depend on *Ne* and the selection coefficient (e.g., see Kimura, 1962). At the same time, favourable alleles may be lost by stochastic sampling, and thereby removed from exposure to selection, or slightly deleterious alleles may go to fixation unless selection against them is strong (e.g., Feng et al., 2019).

4.4 | Are selection scans poorly designed or are adaptive loci simply undetectable?

Our simulations provided estimates of the power and specificity of three R tools for selection analysis (GWDS, PCadapt and OutFLANK) in the context of pairwise source and founder population comparisons in the absence of gene flow, and assuming complete linkage of genotyped SNPs to causative SNPs. We evaluated the performance of each test for various combinations of selection coefficients (*s*) and founder effective population sizes (*Ne*) (Figure 3; Figure S13) in recently diverged populations. The focus of this simulation study differs both in the methodology and aim from earlier simulation studies, which evaluated the performance of selection scans under varying demographic models (De Mita et al., 2013; Lotterhos & Whitlock, 2014; Luu et al., 2017; Narum & Hess, 2011).

Our simulation outcomes are informative for the particular case of heterogeneous selection on standing variation in systems of source–founder populations in the absence of gene flow, and caution should be exercised when extrapolating the results to other demographic scenarios. Furthermore, because our simulation tools simulated allele frequencies on an SNP by SNP basis, they did not allow us to assess the dependency of the performance of selection scans on recombination rates, nor the potential interference with background selection. The results presented in this study hold true only for SNPs which are either causative SNPs themselves or in (near) complete linkage to causative SNPs. A potential shortfall of our simulations is that they did not account for the potential confounding factor of missing data variability across SNPs. Both GWDS and OutFLANK infer an outlier threshold from the observed distribution of SNP by SNP population differentiation scores. It is possible that variation in levels of missing data can affect the shape of this distribution, and as a result bias outlier detection.

Our simulations indicated strong dependency of the performance of selection scans on both *s* and *Ne*, with poor power resulting from low *Ne* and/or low *s*, exacerbated by the sampling effect. For $Ne \leq 50$, the majority of positively selected loci were not detected by any selection test, unless the selection coefficient was high ($s \geq 0.15$). Our simulations suggest that under certain demographic scenarios the software OutFLANK has relatively low power, as reported previously (Bernatchez et al., 2016; Luu et al., 2017). However, our simulations also confirm the claim of OutFLANK developers that for other scenarios OutFLANK has very high specificity without greatly compromising its power (Lotterhos & Whitlock, 2015). The low specificity scores of PCadapt for the pooled approach should not be considered evidence that PCadapt performs poorly, because PCadapt is not designed for the particular purpose to which it was applied in this approach. PCadapt cannot be forced to detect outliers for a prespecified population division, but instead outputs all outliers for all putative population divisions.

Although the original aim of our simulations was to investigate the false discovery rate of selection scans and thereby quantify the probability that the observed outliers in our empirical data sets were not false positives, the simulation results can also be viewed in a different

way. The selection scans OutFLANK and GWDS, presented in this study, are designed to mark SNPs which stand out from the overall distribution. The implicit assumption behind these tests is that SNPs under selection will stand out from the overall distribution. However, this assumption is often violated (Figure S15). Therefore, if assuming that OutFLANK and GWDS correctly mark SNPs which stand out, the selection scan power estimates can be taken as estimates of the efficiency of selection in the context of the underlying demographic scenario. In other words, the obtained power estimates of the selection scans do not in fact reflect the selection scan itself, but rather the outcome of the interplay between selection and genetic drift. Simulated Fdist plots (Figure S15) indeed indicate that the low power and specificity of selection scans in the context of small founder populations probably reflect the confounding effect of drift overriding and obscuring positive selection rather than flawed test designs.

Genetic drift can make selected loci indistinguishable from neutral loci in two ways. Drift can moderate or even counteract selection-driven allele frequency change, and drift affects the backdrop of neutral variation from which selected loci need to stand out in order to be detected by selection scans (Whitlock & Lotterhos, 2015). In small isolated populations the time window in which positively selected loci can stand out from the backdrop of neutral variation (i.e., approach and reach fixation before neutral alleles do so) is limited or near absent (Figure S14). In larger populations, in contrast, neutral alleles take longer to reach fixation, which provides a time window for adaptive loci to stand out. In the presence of gene flow, the allele frequencies in populations are correlated, and F_{ST} values do not converge to 1, extending the time window in which heterogeneous selection can make adaptive loci stand out.

Visual inspection of simulated H_E–F_{ST} distributions confirmed that whether selected loci stand out from neutral loci depends both on the sample size (i.e., number of genotyped individuals per population) and on the long-term effective population size (*Ne*) of the founder population (Figure S14). For Ne =20, the distribution of neutral alleles nearly fills the entire shark fin-shaped H_E–F_{ST} spectrum, obscuring most loci under selection. For $Ne \geq 50$ there is greater opportunity for loci under selection to be detected (Figure S14).

The simulation results presented in this study should therefore not be interpreted to mean that selection scans overlook many events of positive selection. Instead, the low power estimates indicate that, in the context of founder populations, selection is often overruled by drift. Whereas a previous study (Poh et al., 2014) found that SFS-based selection scans are powerless in the context of recently bottlenecked founder populations, our simulations show that F_{ST}-based selection scans (including GWDS) do provide some power. However, we also stress that F_{ST}-based selection scans can detect positive selection events during a restricted time window only. The upper boundary of this time window is roughly 4*Ne* generations after the vicariance event, assuming the absence of gene flow. Outside of this time window, other types of selection scans will prove more effective (Oleksyk et al., 2010).

5 | CONCLUSIONS

We screened for signals of selection in two recently established and heavily bottlenecked reindeer founder populations using a ddRADseq data set, and detected two adjacent SNPs which were marked as outliers by all four selection scans and which potentially result from parallel positive selection. We furthermore presented a new selection scan, called GWDS, well suited to our study system, together with Wright–Fisher model simulations which generated power and specificity estimates of GWDS and of PCadapt and OutFLANK. Our simulations showed that positive selection events in small founder populations are most likely detected by GWDS. We hypothesize that our detected genetic signals of selection correspond to differential survival rates and consequent fitness variation among individuals with and without mandibular swellings resulting from dental disease. This study therefore provides both *in silico* and empirical evidence that although founder bottlenecks restrict adaptive potential, small founder populations can nevertheless start adapting to their novel environment directly following a founder event, and this signal can be detected using genomic data.

ACKNOWLEDGEMENTS

We thank the Whitehead Trust and the British Deer Society for funding, and Michelle Gaither and Kim Andrews for support with laboratory work and analyses. We thank Jennifer Lee, the Government of South Georgia & South Sandwich Islands, members of BSES Expeditions and the British Antarctic Survey for assistance in obtaining samples from South Georgia. We thank Asgrim Opdal, Olaf Odegaard and Filefjell Reinlag for the provision of samples from Norway.

AUTHOR CONTRIBUTIONS

A.R.H. conceived the study and M.d.J. and A.R.H. wrote the paper and all authors read the manuscript and provided comment. M.d.J. undertook data, simulation and laboratory analyses, and developed the selection scan GWDS. F.L. provided fieldwork and some of the DNA extractions.

DATA AVAILABILITY STATEMENT

Sequences associated with RAD analyses are deposited at GenBank under BioProject accession PRJNA700078. RADseq genotype (SNP) files are provided on Dryad at: https://doi.org/10.5061/dryad.rv15dv47h. There are no restrictions on data availability. The selection scan GWDS is implemented in the R package SambaR, which can be downloaded from: https://github.com/mennodejong1986/SambaR

ORCID

Menno J. de Jong https://orcid.org/0000-0003-2131-9048
A. Rus Hoelzel https://orcid.org/0000-0002-7265-4180

REFERENCES

Ahrens, C. W., Rymer, P. D., Stow, A., Bragg, J., Dillon, S., Umbers, K. D. L., & Dudaniec, R. Y. (2018). The search for loci under selection: trends, biases and progress. *Molecular Ecology*, *27*(6), 1342–1356.

Allendorf, F. W., & Lundquist, L. L. (2003). Introduction: Population biology, evolution, and control of invasive species. *Conservation Biology*, *17*, 24–30.

Andrew, S. C., Jensen, H., Hagen, I. J., Lundregan, S., & Griffith, S. C. (2018). Signatures of genetic adaptation to extremely varied Australian environments in introduced European house sparrows. *Molecular Ecology*, *27*, 4542–4555.

Beaumont, M. A. (2005). Adaptation and speciation: What can F(st) tell us? *Trends in Ecology & Evolution*, *20*, 435–440.

Beaumont, M. A., & Nichols, R. A. (1996). Evaluating loci for use in the genetic analysis of population structure. *Proceedings: Biological Sciences*, *263*, 1619–1626.

Bernatchez, S., Laporte, M., Perrier, C., Sirois, P., & Bernatchez, L. (2016). Investigating genomic and phenotypic parallelism between piscivorous and planktivorous lake trout (*Salvelinus namaycush*) ecotypes by means of RADseq and morphometrics analyses. *Molecular Ecology*, *25*, 4773–4792.

Bovine Genome Sequencing and Analysis Consortium, Elsik, C. G., Tellam, R. L., Worley, K. C., Gibbs, R. A., Muzny, D. M., Weinstock, G. M., Adelson, D. L., Eichler, E. E., Elnitski, L., Guigó, R., Hamernik, D. L., Kappes, S. M., Lewin, H. A., Lynn, D. J., Nicholas, F. W., Reymond, A., Rijnkels, M., Skow, L. C., & Zhao, F.-Q. 2009. The genome sequence of taurine cattle: A window to ruminant biology and evolution. *Science (New York, N.Y.)*, *324*, 522–528.

Bradshaw, W. E., & Holzapfel, C. M. (2010). Light, time, and the physiology of biotic response to rapid climate change in animals. *Annual Review of Physiology*, *72*, 147–166.

Brakefield, P. M., & de Jong, P. W. (2011). A steep cline in ladybird melanism has decayed over 25 years: A genetic response to climate change? *Heredity*, *107*, 574–578.

Cammen, K. M., Schultz, T. F., Rosel, P. E., Wells, R. S., & Read, A. J. (2015). Genomewide investigation of adaptation to harmful algal blooms in common bottlenose dolphins (*Tursiops truncates*). *Molecular Ecology*, *24*, 4697–4710.

Carroll, S. P., Hendry, A. P., Reznick, D. N., & Fox, C. W. (2007). Evolution on ecological time-scales. *Functional Ecology*, *21*, 387–393.

Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: An analysis tool set for population genomics. *Molecular Ecology*, *22*, 3124–3140.

Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: Rising to the challenge of larger and richer datasets. *GigaScience*, *4*, 7.

Chen, Z., Farrell, A. P., Matala, A., Hoffman, N., & Narum, S. R. (2018). Physiological and genomic signatures of evolutionary thermal adaptation in redband trout from extreme climates. *Evolutionary Applications*, *11*, 1686–1699.

Cockerham C. C., & Weir B. S. (1987). Correlations, descent measures: drift with migration and mutation. *Proceedings of the National Academy of Sciences*, *84*(23), 8512–8514. http://dx.doi.org/10.1073/pnas.84.23.8512

Colautti, R. I., & Lau, J. A. (2015). Contemporary evolution during invasion: Evidence for differentiation, natural selection, and local adaptation. *Molecular Ecology*, *24*, 1999–2017.

Darcey, J., Horner, K., Walsh, T., Southern, H., Marjanovic, E. J., & Devlin, H. (2013). Tooth loss and osteoporosis: To assess the association between osteoporosis status and tooth number. *British Dental Journal*, *214*, E10.

De Jong, M. J., De Jong, J. F., Hoelzel, A. R., & Janke, A. (2021). SambaR: an R package for fast, easy and reproducible population-genetic analyses of biallelic SNP datasets. *Molecular Ecology Resources*. https://doi.org/10.1111/1755-0998.13339.

De Jong, M., Lovatt, F., & Hoelzel, A. R. (2020). RADseq short read sequences; Genbank. Persistent identifier: PRJNA700078.

De Jong, M., Lovatt, F., & Hoelzel, A. R. (2020). RADseq genotypes; Dryad. Persistent identifier: https://doi.org/10.5061/dryad.rv15dv47h.

De Mita, S., Thuillet, A.-C., Gay, L., Ahmadi, N., Manel, S., Ronfort, J., & Vigouroux, Y. (2013). Detecting selection along environmental gradients: Analysis of eight methods and their effectiveness for outbreeding and selfing populations. *Molecular Ecology*, *22*, 1383–1399.

Duforet-Frebourg, N., Bazin, E., & Blum, M. G. B. (2014). Genome scans for detecting footprints of local adaptation using a Bayesian factor model. *Molecular Biology and Evolution*, *31*, 2483–2495.

Endler, J. (1986). Natural Selection in the Wild.

Feng, S., Fang, Q., Barnett, R., Li, C., Han, S., Kuhlwilm, M., Zhou, L., Pan, H., Deng, Y., Chen, G., Gamauf, A., Woog, F., Prys-Jones, R., Marques-Bonet, T., Gilbert, M. T. P., & Zhang, G. (2019). The genomic footprints of the fall and recovery of the crested Ibis. *Current Biology: CB*, *29*, 340–349.e7.

Foll, M., & Gaggiotti, O. (2008). A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A bayesian perspective. *Genetics*, *180*, 977–993.

Foster, J. B. (1964). Evolution of mammals on Islands. *Nature*, *202*, 234–235.

Frichot, E., & François, O. (2015). LEA: An R package for landscape and ecological association studies. *Methods in Ecology and Evolution*, *6*, 925–929.

Fuller, J., Ferchaud, A.-L., Laporte, M., Le Luyer, J., Davis, T. B., Cote, S. V., & Bernatchez, L. (2020). Absence of founder effect and evidence for adaptive divergence in a recently introduced insular population of white-tailed deer (*Odocoileus virginianus*). *Molecular Ecology*, *29*(1), 86–104.

Funk, W. C., Lovich, R. E., Hohenlohe, P. A., Hofman, C. A., Morrison, S. A., Sillett, T. S., Ghalambor, C. K., Maldonado, J. E., Rick, T. C., Day, M. D., Polato, N. R., Fitzpatrick, S. W., Coonan, T. J., Crooks, K. R., Dillon, A., Garcelon, D. K., King, J. L., Boser, C. L., Gould, N., & Andelt, W. F. (2016). Adaptive divergence despite strong genetic drift: genomic analysis of the evolutionary mechanisms causing genetic differentiation in the island fox (*Urocyon littoralis*). *Molecular Ecology*, *25*(10), 2176–2194.

Guo, Z., Neilson, L. J., Zhong, H., Murray, P. S., Zanivan, S., & Zaidel-Bar, R. (2014). E-cadherin interactome complexity and robustness resolved by quantitative proteomics. *Science Signaling*, *7*, rs7.

Hendrickson, S. L. (2013). A genome wide study of genetic adaptation to high altitude in feral Andean horses of the paramo. *BMC Evolutionary Biology*, *13*, 273.

Hendry, A. P., & Kinnison, M. T. (1999). Perspective: The pace of modern life: measuring rates of contemporary microevolution. *Evolution*, *53*, 1637–1653.

Heppenheimer, E., Brzeski, K. E., Hinton, J. W., Patterson, B. R., Rutlegde, L. Y., DeCandia, A. L., Wheeldon, T., Fain, S. R., Hohenlohe, P. A., Kays, R., White, B. N., Chamberlain, M. J., & vonHoldt, B. M. (2018). High genomic diversity and candidate genes under selection associated with range expansion in eastern coyote (Canis latrans) populations. *Ecology and Evolution*, *8*(12641), 12655.

Hermisson, J., & Pennings, P. S. (2005). Soft sweeps: Molecular population genetics of adaptation from standing genetic variation. *Genetics*, *169*, 2335–2352.

Heymann, R., About, I., Lendahl, U., Franquin, J.-C., Öbrink, B., & Mitsiadis, T. A. (2002). E- and N-cadherin distribution in developing and functional human teeth under normal and pathological conditions. *American Journal of Pathology*, *160*, 2123–2133.

Hill, W. G., & Weir, B. S. (1988). Variances and covariances of squared linkage disequilibria in finite populations. *Theoretical Population Biology*, *33*, 54–78.

Hof, A. E., Campagne, P., Rigden, D. J., Yung, C. J., Lingley, J., Quail, M. A., Hall, N., Darby, A. C., & Saccheri, I. J. (2016). The industrial melanism mutation in British peppered moths is a transposable element. *Nature*, *534*, 102–105.

Johnston, R. F., & Selander, R. K. (1964). House sparrows: Rapid evolution of races in North America. *Science (New York, N.Y.)*, *144*, 548–550.

Johnston, S. E., Huisman, J., Ellis, P. A., & Pemberton, J. M. (2017). A high-density linkage maps reveals sexual dimorphism in recombination landscapes in Red Deer (*Cervus elaphus*). *G3: Genes, Genomes, Genetics*, *7*(8), 2859–2870.

Jombart, T. (2008). adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*, 1403–1405.

Jombart, T., & Ahmed, I. (2011). Adegenet 1.3-1: New tools for the analysis of genome-wide SNP data. *Bioinformatics*, *27*, 3070–3071.

Karell, P., Ahola, K., Karstinen, T., Valkama, J., & Brommer, J. E. (2011). Climate change drives microevolution in a wild bird. *Nature Communications*, *2*, 208.

Kawakami, T., Backstrom, N., Burri, R., Husby, A., Olason, P., Rice, A. M., Alund, M., Qvarnstrom, A., & Ellegren, H. (2014) Estimation of linkage disequilibrium and interspecific gene flow in Ficedula flycatchers by a newly developed 50k single-nucleotide polymorphism array. *Molecular Ecology Resources*, *14*, 1246–1260.

Kimura, M. (1962). On the probability of fixation of mutant genes in a population. *Genetics*, *47*, 713–719.

Kimura, M., & Ohta, T. (1969). The average number of generations until fixation of a mutant gene in a finite population. *Genetics*, *61*, 763–771.

Klein, D. R., Meldgaard, M., & Fancy, S. G. (1987). Factors determining leg length in *Rangifer tarandus*. *Journal of Mammalogy*, *277*, 642–1669.

Lamichhaney, S., Berglund, J., Almén, M. S., Maqbool, K., Grabherr, M., Martinez-Barrio, A., Promerová, M., Rubin, C.-J., Wang, C., Zamani, N., Grant, B. R., Grant, P. R., Webster, M. T., & Andersson, L. (2015). Evolution of Darwin's finches and their beaks revealed by genome sequencing. *Nature*, *518*, 371–375.

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*, 357–359.

Leader-Williams, N. (1980). Population dynamics and mortality of reindeer introduced into South Georgia. *Journal of Wildlife Management*, *44*, 640–657.

Leader-Williams, N. (1982). Relationship between a disease, host density and mortality in a free-living deer population. *Journal of Animal Ecology*, *51*, 235–240.

Leader-Williams, N. (1988). *Reindeer on South Georgia* (p. 319). Cambridge University Press.

Lee, K. M., & Coop, G. (2019). Population genomics perspectives on convergent adaptation. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *374*, 20180236.

Lepais, O., & Weir, J. T. (2014). SimRAD: An R package for simulation-based prediction of the number of loci expected in RADseq and similar genotyping by sequencing approaches. *Molecular Ecology Resources*, *14*, 1314–1321.

Li, C.-Y., Cha, W., Luder, H.-U., Charles, R.-P., McMahon, M., Mitsiadis, T. A., & Klein, O. D. (2012). E-cadherin regulates the behavior and fate of epithelial stem cells and their progeny in the mouse incisor. *Developmental Biology*, *366*, 357–366.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., & Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, *25*, 2078–2079.

Li, Z., Lin, Z., Ba, H., Chen, L., Yang, Y., Wang, K., Qiu, Q., Wang, W., & Li, G. (2017). Draft genome of the reindeer (*Rangifer tarandus*). *GigaScience*, *6*, 1–5.

Lischer, H. E. L., & Excoffier, L. (2012). PGDSpider: An automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics*, *28*, 298–299.

Lomolino, M. V., van der Geer, A. A., Lyras, G. A., Palombo, M. R., Sax, D. F., & Rozzi, R. (2013). Of mice and mammoths: Generality and antiquity of the island rule. *Journal of Biogeography*, *40*, 1427–1439.

Losos, J. B., & Ricklefs, R. E. (2009). Adaptation and diversification on islands. *Nature*, *457*, 830–836.

Lotterhos, K. E., & Whitlock, M. C. (2014). Evaluation of demographic history and neutral parameterization on the performance of FST outlier tests. *Molecular Ecology*, *23*, 2178–2192.

Lotterhos, K. E., & Whitlock, M. C. (2015). The relative power of genome scans to detect local adaptation depends on sampling design and statistical method. *Molecular Ecology*, *24*, 1031–1046.

Lovatt, F. M., & Hoelzel, A. R. (2011). The impact of population bottlenecks on fluctuating asymmetry and morphological variance in two separate populations of reindeer on the island of South Georgia. *Biological Journal of the Linnean Society*, *102*, 798–811.

Lovatt, F. M., & Hoelzel, A. R. (2014). Impact on reindeer (*Rangifer tarandus*) genetic diversity from two parallel population bottlenecks founded from a common source. *Evolutionary Biology*, *41*, 240–250.

Lowry, D. B., Hoban, S., Kelley, J. L., Lotterhos, K. E., Reed, L. K., Antolin, M. F., & Storfer, A. (2016). Breaking RAD: an evaluation of the utility of restriction site-associated DNA sequencing for genome scans of adaptation. *Molecular Ecology Resources*, *17*, 142–152.

Luu, K., Bazin, E., & Blum, M. G. B. (2017). pcadapt: An R package to perform genome scans for selection based on principal component analysis. *Molecular Ecology Resources*, *17*, 67–77.

McVean, G. A. T., & Cardin, N. J. Approximating the coalescent with recombination. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1459), 1387–1393.

Mussmann, S. M., Douglas, M. R., Chafin, T. K., & Douglas, M. E. (2019). BA3-SNPs: Contemporary migration reconfigured in BayesAss for next-generation sequence data. *Methods in Ecology and Evolution*, *10*, 1808–1813.

Narum, S. R., & Hess, J. E. (2011). Comparison of FST outlier tests for SNP loci under selection. *Molecular Ecology Resources*, *11*, 184–194.

Nei, M. (1972). Genetic distance between populations. *American Naturalist*, *106*, 283–292.

Oleksyk, T. K., Smith, M. W., & O'Brien, S. J. (2010). Genome-wide scans for footprints of natural selection. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*, 185–205.

Paradis, E., Claude, J., & Strimmer, K. (2004). APE: analyses of phylogenetics and evolution in R language. *Bioinformatics*, *20*, 289–290.

Paradis, E., & Schliep, K. (2018). ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, *35*, 526–528.

Pembleton, L. W., Cogan, N. O. I., & Forster, J. W. (2013). StAMPP: an R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. *Molecular Ecology Resources*, *13*, 946–952.

Perrier, C., Bourret, V., Kent, M. P., & Bernatchez, L. (2013). Parallel and nonparallel genome-wide divergence among replicate population pairs of freshwater and anadromous Atlantic salmon. *Molecular Ecology*, *22*, 5577–5593.

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS One*, *7*, e37135.

Poh, Y.-P., Domingues, V. S., Hoekstra, H. E., & Jensen, J. D. (2014). On the prospect of identifying adaptive loci in recently bottlenecked populations. *PLoS One*, *9*, e110579.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., de Bakker, P. I. W., Daly, M. J., & Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, *81*, 559–575.

Quinlan, A. R., & Hall, I. M. (2010). BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics*, *26*, 841–842.

Reznick, D. N., & Ghalambor, C. K. (2001). The population ecology of contemporary adaptations: What empirical studies reveal about

the conditions that promote adaptive evolution. *Genetica*, *112–113*, 183–198.

Roesti, M., Gavrilets, S., Hendry, A. P., Salzburger, W., & Berner, D. (2014). The genomic signature of parallel adaptation from shared genetic variation. *Molecular Ecology*, *23*, 3944–3956.

Rozzi, R., & Lomolino, M. V. (2017). Rapid Dwarfing of an Insular Mammal—The Feral Cattle of Amsterdam Island. *Scientific Reports*, *7*, 8820.

Schoener, T. W. (2011). The newest synthesis: understanding the interplay of evolutionary and ecological dynamics. *Science*, *331*, 426–429.

Selvaraj, A., & Prywes, R. (2003). Megakaryoblastic leukemia-1/2, a transcriptional co-activator of serum response factor, is required for skeletal myogenic differentiation. *Journal of Biological Chemistry*, *278*, 41977–41987.

Shultz, A. J., Baker, A. J., Hill, G. E., Nolan, P. M., & Edwards, S. V. (2016). SNPs across time and space: population genomic signatures of founder events and epizootics in the House Finch (*Haemorhous mexicanus*). *Ecology and Evolution*, *6*, 7475–7489.

Swärd, K., Stenkula, K. G., Rippe, C., Alajbegovic, A., Gomez, M. F., & Albinsson, S. (2016). Emerging roles of the myocardin family of proteins in lipid and glucose metabolism. *Journal of Physiology*, *594*, 4741–4752.

Takeuchi, T., Masaoka, T., Aoki, H., Koyanagi, R., Fujie, M., & Satoh, N. (2019). Divergent northern and southern populations and demographic history of the pearl oyster in the western Pacific revealed with genomic SNPs, https://doi.org/10.1111/eva.12905.

Templeton, A. R. (2008). The reality and importance of founder speciation in evolution. *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology*, *30*, 470–479.

Vandepitte, K., de Meyer, T., Helsen, K., van Acker, K., Roldán-Ruiz, I., Mergeay, J., & Honnay, O. (2014). Rapid genetic adaptation precedes the spread of an exotic plant species. *Molecular Ecology*, *23*, 2157–2164.

Weigand, H., & Leese, F. (2018). Detecting signatures of positive selection in non-model species using genomic data. *Zoological Journal of the Linnean Society*, *184*, 528–583.

Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution*, *38*, 1358–1370.

Whitlock, M. C., & Lotterhos, K. (2014). OutFLANK: Fst outliers with trimming.

Whitlock, M. C., & Lotterhos, K. E. (2015). Reliable detection of loci responsible for local adaptation: Inference of a null model through trimming the distribution of F(ST). *American Naturalist*, *186*(Suppl 1), S24–36.

Willi, Y., van Buskirk, J., & Hoffmann, A. A. (2006). Limits to the adaptive potential of small populations. *Annual Review of Ecology, Evolution, and Systematics*, *37*, 433–458.

Zimin, A. V., Delcher, A. L., Florea, L., Kelley, D. R., Schatz, M. C., Puiu, D., Hanrahan, F., Pertea, G., Van Tassell, C. P., Sonstegard, T. S., Marçais, G., Roberts, M., Subramanian, P., Yorke, J. A., & Salzberg, S. L. (2009). A whole-genome assembly of the domestic cow, *Bos taurus*. *Genome Biology*, *10*, R42.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Jong MJ, Lovatt F, Hoelzel AR. Detecting genetic signals of selection in heavily bottlenecked reindeer populations by comparing parallel founder events. *Mol Ecol*. 2021;30:1642–1658. https://doi.org/10.1111/mec.15837